MIMO Multiple Access Channel with an Arbitrarily Varying Eavesdropper

Xiang He, Ashish Khisti*, Aylin Yener†
*Dept. of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada
† Electrical Engineering Department, The Pennsylvania State University, University Park, PA 16802
xianghe@microsoft.com, akhisti@comm.utoronto.ca, yener@psu.edu

***Abstract*—A two-transmitter Gaussian multiple access wiretap channel with multiple antennas at each of the nodes is investigated. The channel matrices at the legitimate terminals are fixed and revealed to all the terminals, whereas the channel matrix of the eavesdropper is arbitrarily varying and only known to the eavesdropper. The secrecy degrees of freedom (s.d.o.f.) region under a strong secrecy constraint is characterized. A transmission scheme that orthogonalizes the transmit signals of the two users at the intended receiver and uses a single-user wiretap code is shown to be sufficient to achieve the s.d.o.f. region. The converse involves establishing an upper bound on a weighted-sum-rate expression. This is accomplished by using induction, where at each step one combines the secrecy and multiple-access constraints associated with an adversary eavesdropping a carefully selected group of sub-channels.**

I. Introduction

Information theoretic security was first introduced by Shannon in [1], which studied the problem of transmitting confidential information in a communication system in the presence of an eavesdropper with unbounded computational power. Since then, an extensive body of work has been devoted to studying this problem for different network models by deriving fundamental transmission rate limits [2]–[4] and designing low-complexity schemes to approach these limits in practice [5], [6].

Secure communication using multiple antennas was extensively studied as well, see e.g., [7]–[15]. These works investigated efficient signaling mechanisms using the spatial degrees of freedom provided by multiple antennas to limit an eavesdropper's ability to decode information. The underlying information theoretic problem, the multi-antenna wiretap channel, was studied and the associated secrecy capacity was identified. We note that these works assumed that the eavesdropper's channel state information is available either completely or partially, although such an assumption may not be justified in practice.

As a more pessimistic but stronger assumption, references [16]–[18] study secrecy capacity when the eavesdropper channel is arbitrarily varying and its channel states are known to the eavesdropper only. Reference [17] studies the single-user Gaussian multi-input-multi-output (MIMO) wiretap channel and characterizes the secrecy degrees of freedom (s.d.o.f.). The same paper extended the single user analysis to the two user Gaussian MIMO multiple access (MIMO-MAC) channel. This was possible only when all the legitimate terminals have equal number of antennas, leaving the MIMO-MAC with arbitrary number of antennas at the terminals an open problem.

Our main contribution is to fully characterize the s.d.o.f. region of the two-transmitter MIMO MAC channel when the eavesdropper channel is arbitrarily varying. We show that the s.d.o.f. region can be achieved by a scheme that orthogonalizes the transmit signals of the two users at the intended receiver. Moreover, it suffices to use a single-user wiretap channel code [17] and no coordination between the users is necessary except for synchronization and sharing the transmit dimensions. To establish the optimality of this scheme, our converse proof decomposes the MIMO MAC channel into a set of parallel and independent channels using the generalized singular value decomposition (GSVD). A set of eavesdroppers, each monitoring a subset of links, is selected using an induction procedure and the resulting secrecy constraints are combined to obtain an upper bound on a weighted sum-rate expression. The outer bound matches the achievable rate in terms of the s.d.o.f. region, thus settling the open problem raised in [17] for the case of two transmitters.

Interestingly, the s.d.o.f. region remains open for this model when the eavesdropper channel is perfectly known to all terminals. A significant body of literature already exists on this problem, see e.g., [19]–[22]. If the channel model has real inputs and outputs, Gaussian signaling is in general suboptimal and user cooperating strategies as well as signal alignment techniques are necessary [23]. In [24] it is established that s.d.o.f. of $1/2$ is achievable using real interference alignment for almost all configurations of channel gains. If the channel model has complex inputs and outputs, it is shown in [25, Section 5.16] that in general s.d.o.f. of $1/2$ is achievable using asymmetric Gaussian signaling. In contrast, the best known upper bound on the s.d.o.f. of individual rates is $2/3$ for both cases, established in [25, Section 5.5].

The remainder of this paper is organized as follows. In Section II, we describe the system model. The main result is stated as Theorem 1 in Section IV. The proof of the theorem is divided into two parts. First, we establish the result for the case of parallel channels in Section V. Subsequently, in Section VI we establish the result for the general case by decomposing the MIMO-MAC channel into a set of independent parallel

This work was presented in part at the 49th Annual Allerton Conference on Communication, Control, and Computing, September, 2011. This work is supported in part by NSF Grant 0964362. A. Khisti's work was supported by an NSERC Discovery Grant. The ordering of authors is alphabetical.

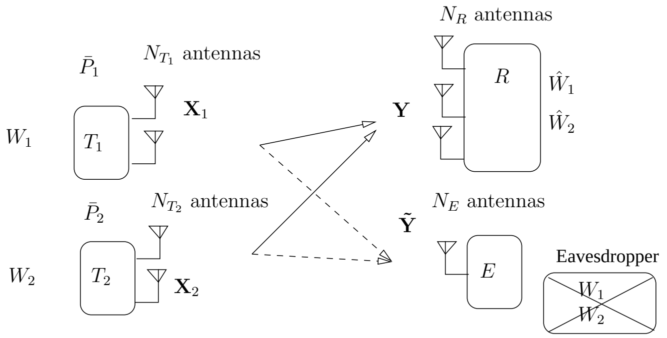

Fig. 1. The MIMO MAC wiretap channel where $N_{T_1} = N_{T_2} = 2$, $N_R = 3$, $N_E = 1$.

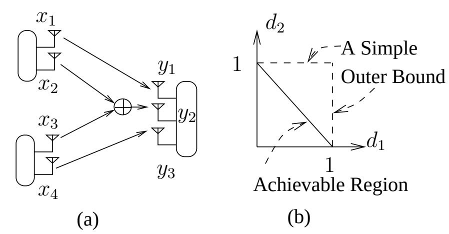

Fig. 2. (a) A special case of MIMO MAC wiretap channel where $N_{T_1} = N_{T_2} = 2$, $N_R = 3$, $N_E = 1$, (b) Comparison between achievable s.d.o.f. region and a simple outer bound derived by considering one eavesdropper at a time.

channels. Such a reduction is used both in the proof of the converse as well as the coding scheme. Section VII concludes the paper.

We use the following notation throughout the paper: For a set $\mathcal{A}$, $V_{i,\mathcal{A}}$ and $V_{\mathcal{A}}$ denote the set of random variables $\{V_{i,j}, j \in \mathcal{A}\}$ and $\{V_j, j \in \mathcal{A}\}$ respectively. $\{\delta_n\}$ denotes a non-negative sequence of n that converges to 0 when n goes to ∞. We use bold upper-case font for matrices and vectors and lower-case font for scalars. The distinction between matrices and vectors will be clear from the context. For a set $\mathcal{A}$, $|\mathcal{A}|$ denotes its cardinality and a short hand notation x^n is used for the sequence $\{x_1, x_2, \ldots, x_n\}$. ϕ denotes the empty set.

II. System Model

As shown in Figure 1, we consider a discrete-time channel model where two transmitters communicate with one receiver in the presence of an eavesdropper. We assume transmitter i has N_{T_i} antennas, $i = 1, 2$, the legitimate receiver has N_R antennas whereas the eavesdropper has N_E antennas. The channel model is given by

$$\mathbf{Y}(i) = \sum_{k=1}^{2} \mathbf{H}_k \mathbf{X}_k(i) + \mathbf{Z}(i) \quad (1)$$

$$\tilde{\mathbf{Y}}(i) = \sum_{k=1}^{2} \tilde{\mathbf{H}}_k(i) \mathbf{X}_k(i) \quad (2)$$

where $i \in \{1, \ldots, n\}$ denotes the time-index, $\mathbf{H}_k, k = 1, 2$, are channel matrices and $\mathbf{Z}$ is the additive Gaussian noise observed by the intended receiver, which is composed of independent rotationally invariant complex Gaussian random variables with zero mean and unit variance. The sequence of eavesdropper channel matrices $\{\tilde{\mathbf{H}}_k(i), k = 1, 2\}$, is an arbitrary sequence of length n and only revealed to the eavesdropper. In contrast, $\mathbf{H}_k, k = 1, 2$ are revealed to both the legitimate parties *and* the eavesdropper(s). We assume N_E, the number of eavesdropper antennas, is known to the legitimate parties and the eavesdropper.

User k, $k = 1, 2$, wishes to transmit a confidential message W_k, $k = 1, 2$, to the receiver over n channel uses, while both messages, W_1 and W_2, must be kept confidential from the eavesdropper. We use γ to index a specific sequence of $\{\tilde{\mathbf{H}}_k(i), k = 1, 2\}$ over n channel uses and use $\tilde{\mathbf{Y}}_\gamma^n$ to represent the corresponding channel outputs for $\tilde{\mathbf{Y}}^n$. The *strong* secrecy constraint is [17]:

$$\lim_{n \to \infty} I\left(W_1, W_2; \tilde{\mathbf{Y}}_\gamma^n\right) = 0, \quad \forall \gamma \quad (3)$$

where the convergence must be uniform over γ. The average power constraints for the two users are given by

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} |\mathbf{X}_k(i)|^2 \leq \bar{P}_k, \quad k = 1, 2. \quad (4)$$

The secrecy rate for user k, $R_{s,k}$, is defined as

$$R_{s,k} = \lim_{n \to \infty} \frac{1}{n} H(W_k), \quad k = 1, 2. \quad (5)$$

such that W_k can be reliably decoded by the receiver, and (3) and (4) are satisfied.

We define the secrecy degrees of freedom as:

$$\left\{(d_1, d_2) : d_k = \limsup_{\bar{P}_1 = \bar{P}_2 = \bar{P} \to \infty} \frac{R_{s,k}}{\log_2 \bar{P}}, \quad k = 1, 2\right\} \quad (6)$$

III. Motivation

Before stating the main result, we illustrate the main difficulty in characterizing the s.d.o.f. region through a simple example. As illustrated in Figure 2(a), in this example, each transmitter has 2 antennas and the intended receiver has 3 antennas, while the eavesdropper has only 1 antenna. Let x_1, x_2, x_3, x_4 denote the transmitted signals from the two users and y_1, y_2, y_3 denote the signals observed by the intended the receiver. And the main channel is given by

$$y_1 = x_1 + z_1, \quad y_3 = x_4 + z_3 \quad (7)$$

$$y_2 = x_2 + x_3 + z_2 \quad (8)$$

where $z_i, i = 1, 2, 3$ denote additive channel noise. As shown in [17], a secrecy degree of freedom $\min(N_{T_k}, N_R) - N_E = 1$ is achievable for a user if the other user remains silent. Time sharing between these two users lead to the following achievable s.d.o.f. region:

$$d_1 + d_2 \leq 1, \quad d_k \geq 0, \quad k = 1, 2 \quad (9)$$

For the converse, we begin by considering a simple upper bound, which reduces each channel to a single-user MIMO wiretap channel. First, by revealing the signals transmitted by user 2 to the intended receiver and assuming that the eavesdropper monitors either x_1 or x_2 we have that $d_1 \leq 1$.

Similarly we argue that $d_2 \leq 1$. To obtain an upper bound on the sum-rate we let the two transmitters to cooperate and reduce the system to a 3×3 MIMO link. The s.d.o.f. of this channel [17] yields $d_1 + d_2 \leq 2$. This outer bound, illustrated in Figure 2(b), does not match with the achievable region given by (9).

As we shall show in Theorem 1, (9) is indeed the s.d.o.f. capacity region and hence a new converse is necessary to prove this result. Our key observation is that the above upper bound only considers one eavesdropper at a time in deriving each of the three bounds. For example, when deriving $d_1 \leq 1$, we assume there is only one eavesdropper which is monitoring either x_1 or x_2. When deriving $d_2 \leq 1$, we assume there is only one eavesdropper which is monitoring either x_3 or x_4. Similarly when deriving $d_1 + d_2 \leq 2$ we again assume that there is one eavesdropper on either of the links. As we shall discuss below, a tighter upper bound is possible to find if we consider the simultaneous effect of two eavesdroppers.

In our system model, there are infinitely many possible eavesdroppers, each corresponding to a different channel state sequence. The challenge is to find out a finite number of eavesdroppers, whose joint effect leads to a tight converse. Our choice of eavesdroppers is based on the following intuition: When an eavesdropper chooses which links to monitor, it should give precedence to those links over which only one user can transmit. This is because these links are the major contributor to the sum s.d.o.f. $d_1 + d_2$ since they are dedicated links to a certain user. Based on this intuition, we consider the following two eavesdroppers: one monitors y_1 for W_1 and the other monitors y_3 for W_2. As we shall show later in Lemma 1, the first eavesdropper implies the following upper bound on R_1:

$$n(R_1 - \delta_n) \leq I\left(x_2^n; y_2^n | y_1^n, x_{\{3,4\}}^n\right) \tag{10}$$

and the second eavesdropper implies the following upper bound on R_2:

$$n(R_2 - \delta_n) \leq I\left(y_1^n, x_{\{3,4\}}^n; y_2^n\right) \tag{11}$$

Their joint effect can be captured by adding (10) and (11) [26], which lead to:

$$n(R_1 + R_2 - 2\delta_n) \leq I\left(x_2^n, y_1^n, x_{\{3,4\}}^n; y_2^n\right) \tag{12}$$

Since there is only one term, which is y_2^n, at the right side of the mutual information $I\left(x_2^n, y_1^n, x_{\{3,4\}}^n; y_2^n\right)$, we observe the sum s.d.o.f. can not exceed 1, thereby justifying that (9) is indeed the largest possible s.d.o.f. region for Figure 2(a).

As captured by (10) and (11), a simultaneous selection of two different eavesdroppers for the two users reduces the effective signal dimension at the receiver from three to one, thus leading to a tighter converse. As we shall show later in Section V-C, in generalizing this example we are required to systematically select a sequence of eavesdroppers using an induction procedure.

IV. Main Result

In this section, we state the main result of this work. To express our result, we define r_t as the rank of $\mathbf{H}_t, t = 1, 2$ and r_0 as the rank of $[\ \mathbf{H}_1 \mid \mathbf{H}_2\ ]$. We will refer to r_t as the number of transmit dimensions at user $t = 1, 2$ and r_0 as the number of dimensions at the receiver.

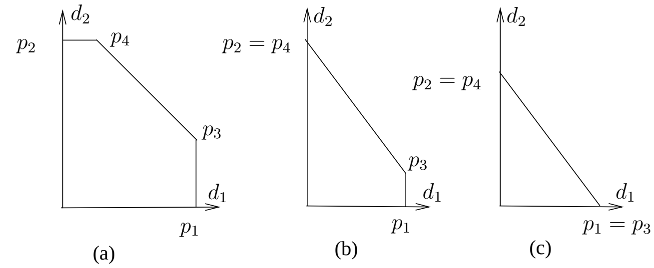

Fig. 3. The secrecy degrees of freedom (s.d.o.f.) region in Theorem 1: (a) $0 \leq N_E \leq \min\{r_0 - r_1, r_0 - r_2\}$, (b) $\min\{r_0 - r_1, r_0 - r_2\} \leq N_E \leq \max\{r_0 - r_1, r_0 - r_2\}$, (c) $\max\{r_0 - r_1, r_0 - r_2\} \leq N_E$

Theorem 1: The secrecy degrees of freedom region of the MIMO multiple access channel with arbitrarily varying eavesdropper channel is given by the convex hull of the following five points of (d_1, d_2):

$$p_0 = (0, 0) \tag{13}$$

$$p_1 = \left([r_1 - N_E]^+, 0\right) \tag{14}$$

$$p_2 = \left(0, [r_2 - N_E]^+\right) \tag{15}$$

$$p_3 = \left([r_1 - N_E]^+, [r_0 - r_1 - N_E]^+\right) \tag{16}$$

$$p_4 = \left([r_0 - r_2 - N_E]^+, [r_2 - N_E]^+\right) \tag{17}$$

where we use $[x]^+ \stackrel{\Delta}{=} \max\{x, 0\}$.

Fig. 3 illustrates the structure of the s.d.o.f. region as a function of the number of eavesdropping antennas. In Fig. 3 (a) we have $N_E \leq \min(r_0 - r_1, r_0 - r_2)$. In this case the s.d.o.f. region is a polymatroid (see e.g., [27, Definition 3.1]) described by $d_i \leq r_i - N_E$ and $d_1 + d_2 \leq r_0 - 2N_E$. Fig. 3 (b) illustrates the shape of the s.d.o.f. region when $\min\{r_0 - r_1, r_0 - r_2\} \leq N_E \leq \max\{r_0 - r_1, r_0 - r_2\}$. In Fig. 3 (b), without loss of generality, we assume $r_1 < r_2$ and the s.d.o.f. region is bounded by the lines $d_i \geq 0$, $d_1 \leq r_1 - N_E$ and

$$\begin{aligned}(r_1 + r_2 - r_0)d_1 + (r_1 - N_E)d_2 \\ \leq (r_1 - N_E) \times (r_2 - N_E).\end{aligned} \tag{18}$$

When $\min(r_1, r_2) > N_E \geq \max(r_0 - r_1, r_0 - r_2)$, the s.d.o.f. region, as illustrated in Fig. 3 (c) is bounded by $d_i \geq 0$ and the line

$$\frac{d_1}{r_1 - N_E} + \frac{d_2}{r_2 - N_E} \leq 1. \tag{19}$$

The s.d.o.f. region in Theorem 1 allows the following simple interpretation: The region can be expressed as a convex hull of a set of rectangles shown by Figure 4 (illustrated for Figure 3 (a)). Each rectangle is parameterized by the dimensions of the subspace occupied by the transmission signals from the two users, denoted by (t_1, t_2), where t_i indicates the dimension of user i, $i = 1, 2$. Then in order for the signals from both transmitters to be received reliably

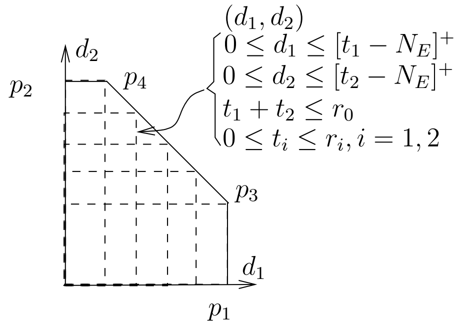

Fig. 4. Interpretation of the s.d.o.f. region as a convex hull of rectangles: $(d_1, d_2) : 0 \le d_i \le [t_i - N_E]^+, i = 1, 2$, where t_i is the number of degrees of freedom occupied by user i. To achieve reliable transmission, we must have (20) and (21).

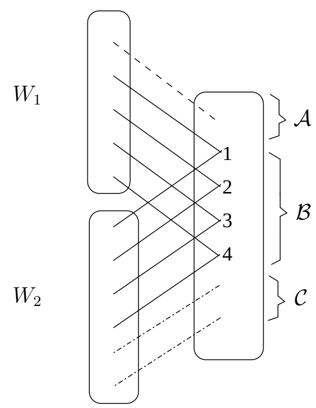

Fig. 5. Definition of the set $\mathcal{A}, \mathcal{B}, \mathcal{C}$, where $|\mathcal{B}| = 4$.

by the receiver, we must have

$$t_1 + t_2 \le r_0 \tag{20}$$
$$0 \le t_i \le r_i, i = 1, 2 \tag{21}$$

Each user then transmits confidential messages with $0 \le d_i \le [t_i - N_E]^+$ over the available t_i dimensions, where the $-N_E$ term is an effect of the secrecy constraint (3).

It is clear that p_3, p_4 given by (16) and (17) are in one of these rectangles. Hence the convex hull of these rectangles yields the s.d.o.f. region stated in Theorem 1.

V. Proof for the Parallel Channel Model

In this section, we establish Theorem 1 for the case of parallel channels. As illustrated in Fig. 5, the receiver observes

$$y_i = x_{1i} + z_i, \quad i \in \mathcal{A}, \tag{22}$$
$$y_i = x_{1i} + x_{2i} + z_i, \quad i \in \mathcal{B}, \tag{23}$$
$$y_i = x_{2i} + z_i, \quad i \in \mathcal{C}, \tag{24}$$

where the noise random variables across the sub-channels are independent and each is distributed according to $\mathcal{CN}(0, 1)$ and $\{x_{1i}\}_{i \in \mathcal{A} \cup \mathcal{B}}$ and $\{x_{2i}\}_{i \in \mathcal{B} \cup \mathcal{C}}$ denote the transmit symbols of user 1 and user 2 respectively.

The parallel channel model is a special case of (1) with

$$\mathbf{H}_1 = \begin{bmatrix} \mathbf{I}_{|\mathcal{A}|} & & \\ & \mathbf{I}_{|\mathcal{B}|} & \\ & & \mathbf{O}_{|\mathcal{C}|} \end{bmatrix}, \ \mathbf{H}_2 = \begin{bmatrix} \mathbf{O}_{|\mathcal{A}|} & & \\ & \mathbf{I}_{|\mathcal{B}|} & \\ & & \mathbf{I}_{|\mathcal{C}|} \end{bmatrix}, \tag{25}$$

where $\mathbf{I}_{|\mathcal{A}|}$, $\mathbf{I}_{|\mathcal{B}|}$ and $\mathbf{I}_{|\mathcal{C}|}$ denote the identity matrices of size $|\mathcal{A}|$, $|\mathcal{B}|$ and $|\mathcal{C}|$ respectively, and $\mathbf{O}_{|\mathcal{A}|}$ and $\mathbf{O}_{|\mathcal{B}|}$ denote the matrices, all of whose entries are zeros. Note that we do not make any assumption on the eavesdropper's channel model (2).

A. Achievability

It suffices to establish the achievability of points p_3 and p_4 in (16) and (17) respectively. The rest of the region follows through time-sharing between these points. Note that for the proposed parallel channel model

$$p_3 = \left([|\mathcal{A}| + |\mathcal{B}| - N_E]^+, [|\mathcal{C}| - N_E]^+\right) \tag{26}$$
$$p_4 = \left([|\mathcal{A}| - N_E]^+, [|\mathcal{B}| + |\mathcal{C}| - N_E]^+\right) \tag{27}$$

To prove the achievability of p_3 we restrict user 2 to transmit only on the last $|\mathcal{C}|$ components of in (24) and allow user 1 to transmit over all of the components of $\mathcal{A} \cup \mathcal{B}$ in (22) and (23). Note that in this case, the signals of these two users do not interfere with each other at the intended receiver. From [17], user 1 can transmit W_1 such that $d_1 = [|\mathcal{A}| + |\mathcal{B}| - N_E]^+$ and

$$\lim_{n \to \infty} I(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n) = 0 \tag{28}$$

and user 2 can transmit W_2 such that $d_2 = [|\mathcal{C}| - N_E]^+$ and

$$\lim_{n \to \infty} I(W_2; \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n) = 0 \tag{29}$$

where we use $\tilde{\mathbf{H}}_k^n \mathbf{X}_k^n$ to denote the sequence $\{\tilde{\mathbf{H}}_k(i) \mathbf{X}_k(i), i = 1, ..., n\}$. Furthermore since $(W_1, \mathbf{X}_1^n$ is independent of $(W_2, \mathbf{X}_2^n)$ we have that

$$\lim_{n \to \infty} I(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n) = 0 \tag{30}$$
$$\lim_{n \to \infty} I(W_2; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n) = 0 \tag{31}$$

which imply:

$$I\left(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n | W_2\right)$$
$$\le I\left(W_1; W_2, \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) \tag{32}$$
$$= I\left(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) + I\left(W_1; W_2 | \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) \tag{33}$$
$$\le I\left(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) + I\left(W_1, \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n; W_2, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) \tag{34}$$
$$= I\left(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n\right) \tag{35}$$

where the last step follows from the fact that $(W_2, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n)$ is independent from $(W_1, \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n)$. Therefore (28) implies

$$\lim_{n \to \infty} I\left(W_1; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n | W_2\right) = 0. \tag{36}$$

Adding (36) and (31), we obtain

$$\lim_{n\to\infty} I\left(W_1, W_2; \tilde{\mathbf{H}}_1^n \mathbf{X}_1^n, \tilde{\mathbf{H}}_2^n \mathbf{X}_2^n\right) = 0 \tag{37}$$

and the secrecy constraint (3) follows from the data-processing inequality. Also, since the convergence in n in (30) and (31) is uniform [17], the convergence in (37) and hence in (3) is uniform as well. Hence we have proved the point p_3 is achievable.

The achievability of p_4 is proved by repeating the argument above by exchanging user 1 with user 2.

Remark 1: As is evident from (37), the secrecy guarantee achieved by one user is not affected by the transmission strategy of the other user. □

B. Converse : $N_E \leq \min(|\mathcal{A}|, |\mathcal{C}|)$

We need to show that the s.d.o.f. region is contained within

$$d_1 \leq |\mathcal{A}| + |\mathcal{B}| - N_E \tag{38}$$

$$d_2 \leq |\mathcal{C}| + |\mathcal{B}| - N_E \tag{39}$$

$$d_1 + d_2 \leq |\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| - 2N_E \tag{40}$$

Since (38) and (39) directly follow from the single user case in [17], we only need to show (40).

Let $\mathcal{E}_k$ be the set of links such that an eavesdropper is monitoring for W_k, $k = 1, 2$. $|\mathcal{E}_1| = |\mathcal{E}_2| = N_E$. $\mathcal{A} \supseteq \mathcal{E}_1$, $\mathcal{C} \supseteq \mathcal{E}_2$. We establish the following upper bound on the achievable rate pairs.

Lemma 1:

$$n(R_{s,1} - \delta_n) \leq I\left(X^n_{1,\mathcal{A}\backslash\mathcal{E}_1}; Y^n_{\mathcal{A}\backslash\mathcal{E}_1}\right) + I\left(X^n_{1,\mathcal{B}}; Y^n_{\mathcal{B}} | M\right) \tag{41}$$

$$n(R_{s,2} - \delta_n) \leq I(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}) + I\left(M; Y^n_{\mathcal{B}}\right) \tag{42}$$

where $M = \left(Y^n_{1,\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right)$.

Proof: The proof is provided in Appendix A. ■

The proof is completed upon adding (41) and (42) so that

$$\begin{aligned} &n(R_{s,1} + R_{s,2} - 2\delta_n) \\ &\leq (X^n_{1,\mathcal{A}\backslash\mathcal{E}_1}; Y^n_{1,\mathcal{A}\backslash\mathcal{E}_1}) + I(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{2,\mathcal{C}\backslash\mathcal{E}_2}) \\ &\quad + I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{B}}) \end{aligned} \tag{43}$$

and using

$$d\left(\frac{1}{n} I(X^n_{\mathcal{A}\backslash\mathcal{E}_1}; Y^n_{\mathcal{A}\backslash\mathcal{E}_1})\right) \leq |\mathcal{A}| - N_E \tag{44}$$

$$d\left(\frac{1}{n} I(X^n_{\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{\mathcal{C}\backslash\mathcal{E}_2})\right) \leq |\mathcal{C}| - N_E \tag{45}$$

$$d\left(\frac{1}{n} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{B}})\right) \leq |\mathcal{B}| \tag{46}$$

where $d(x) \triangleq \lim_{P\to\infty} \frac{x(P)}{\log_2 P}$ characterizes the pre-log scaling of x with respect to P.

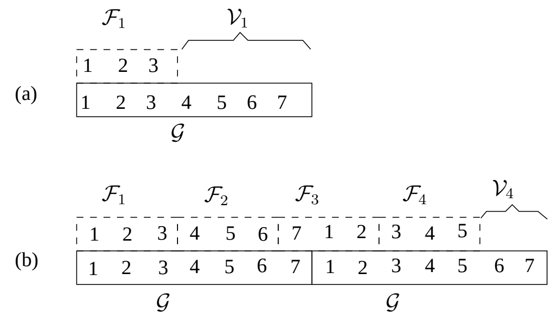

Fig. 6. The set $\mathcal{F}_k$, $\mathcal{G}$, and $\mathcal{V}_k$ when $|\mathcal{F}| = 3$, $|\mathcal{G}| = 7$ and $|\mathcal{B}| = 8$. (a) Case I, $i = 1$, $c_1 = 1$. (b) Case II, $i = 4$, $\mathcal{H}_5 = \{1\}$, $\mathcal{F}_5 = \{6, 7, 1\}$, $\mathcal{V}_5 = \{2, 3, 4, 5, 6, 7\}$, $c_4 = 2$, $c_5 = 3$.

C. Converse : $N_E > \max(|\mathcal{A}|, |\mathcal{C}|)$

Without loss of generality, we assume $|\mathcal{C}| \geq |\mathcal{A}|$. Let $\mathcal{E}_k$ be the set of links such that an eavesdropper is monitoring for W_k, $k = 1, 2$. Let $|\mathcal{E}_1| = |\mathcal{E}_2| = N_E$, $\mathcal{A} \subset \mathcal{E}_1$, and $\mathcal{C} \subset \mathcal{E}_2$.

Define the set $\mathcal{F}, \mathcal{G}$ such that $\mathcal{F} = \mathcal{B}\backslash\mathcal{E}_1$, $\mathcal{G} = \mathcal{B}\backslash\mathcal{E}_2$. Since $|\mathcal{C}| \geq |\mathcal{A}|$, we have $|\mathcal{G}| \geq |\mathcal{F}|$.

Then Theorem 1 reduces to $d_k \geq 0, k = 1, 2$ and

$$|\mathcal{G}| d_1 + |\mathcal{F}| d_2 \leq |\mathcal{F}| \times |\mathcal{G}| \tag{47}$$

which we now show. We first introduce the following lemma:

Lemma 2: For any choice of $\mathcal{F} \subseteq \mathcal{B}$ and $\mathcal{G} \subseteq \mathcal{B}$ with appropriate cardinalities the rates $R_{s,1}$ and $R_{s,2}$ are upper bounded by

$$n(R_{s,1} - \delta_n) \leq I\left(X^n_{1,\mathcal{F}}; Y^n_{\mathcal{F}} | M, X^n_{1,\mathcal{B}\backslash\mathcal{F}}\right) \tag{48}$$

$$n(R_{s,2} - \delta_n) \leq I\left(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{G}}\right) \tag{49}$$

where $M = \left\{Y^n_{1,\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right\}$.

Proof: The proof is provided in Appendix B. ■

For the remainder of the proof we assume without loss of generality that $\mathcal{B} = \{1, \ldots, |\mathcal{B}|\}$. We fix $\mathcal{G} = \{1, \ldots, |\mathcal{G}|\}$ while choosing $|\mathcal{G}|$ different sets of $|\mathcal{F}|$ elements: $\mathcal{F}_1, \ldots, \mathcal{F}_{|\mathcal{G}|}$, the sets $\mathcal{V}_0, \ldots, \mathcal{V}_{|\mathcal{G}|}$ and a sequence of c_i in the following recursive manner.

Definition 1: Let $\mathcal{V}_0 = \mathcal{G}$, $c_0 = 1$. For $i \geq 1$ recursively construct $\mathcal{F}_i$ as follows.

1) **Case I:** $|\mathcal{V}_{i-1}| \geq |\mathcal{F}|$
Let $\mathcal{F}_i = \{\mathcal{V}_{i-1}(1), \ldots, \mathcal{V}_{i-1}(|F|)\}$, where $\mathcal{V}_{i-1}(k)$ denotes the kth smallest element in $\mathcal{V}_{i-1}$. Let $\mathcal{V}_i = \mathcal{V}_{i-1}\backslash\mathcal{F}_i$, and $c_i = c_{i-1}$. This case is illustrated in Figure 6(a) for $i = 1$.
2) **Case II:** $|\mathcal{V}_{i-1}| < |\mathcal{F}|$
Let $\mathcal{F}_i = \mathcal{V}_{i-1} \cup \mathcal{H}_i$, and $\mathcal{V}_i = \mathcal{G}\backslash\mathcal{H}_i$, and $c_i = c_{i-1} + 1$, where $\mathcal{H}_i = \{1, 2, \ldots, |\mathcal{F}| - |\mathcal{V}_{i-1}|\}$. This case is illustrated in Figure 6(b) for $i = 4$.

To interpret the above construction, we note that the set $\mathcal{G}$ is a row-vector with $|\mathcal{G}|$ elements and let $\mathcal{G}^{\otimes}$ be obtained by concatenating $|\mathcal{F}|$ identical copies of the $\mathcal{G}$ vector i.e.,

$$\mathcal{G}^{\otimes} = \underbrace{[\mathcal{G} \mid \mathcal{G} \mid \ldots \mathcal{G}]}_{|\mathcal{F}| \text{ copies}} \tag{50}$$

As shown in Figure 6, by our construction, the vector $\mathcal{F}_1$ spans the first $|\mathcal{F}|$ elements of $\mathcal{G}^{\otimes}$, the vector $\mathcal{F}_2$ spans the next $|\mathcal{F}|$ elements of $\mathcal{G}^{\otimes}$ etc. The constant c_i denotes the index number of copies of the $\mathcal{G}$ vector necessary to cover $\mathcal{F}_i$.

When $i = |\mathcal{G}|$ the row-vector $\mathcal{F}_i$ terminates exactly at the end of the last $\mathcal{G}$ vector in $\mathcal{G}^{\otimes}$. Hence,

$$c_{|\mathcal{G}|} = |\mathcal{F}|, \qquad \mathcal{V}_{|\mathcal{G}|} = \phi. \tag{51}$$

By going through the above recursive procedure and invoking Lemma 2 repeatedly, each time by setting $\mathcal{F}$ in (48) and (49) to be $\mathcal{F}_i$, we establish the following upper bound on the rate region.

Lemma 3: For each $i = 0, 1, \ldots, |\mathcal{G}|$ and the set of channels $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{|\mathcal{G}|}$ defined in Def. 1, the rate pair $(R_{s,1}, R_{s,2})$ satisfies the following upper bound

$$\begin{aligned} & i \cdot n(R_{s,1} - \delta_n) + c_i \cdot n(R_{s,2} - \delta_n) \\ & \le \sum_{j=1}^{i} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_i}). \end{aligned} \tag{52}$$

Before providing a proof, we note that (47) follows from (52) as described below. Evaluating (52) with $i = |\mathcal{G}|$, using (51) and letting $\tilde{R}_{s,i} = R_{s,i} - \delta_n$,

$$\begin{aligned} n|\mathcal{G}|\tilde{R}_{s,1} + n|\mathcal{F}|\tilde{R}_{s,2} &\le \sum_{j=1}^{|\mathcal{G}|} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) && (53) \\ &= \sum_{j=1}^{|\mathcal{G}|} \left\{ h(Y^n_{\mathcal{F}_j}) - h(Y^n_{\mathcal{F}_j} | M, X^n_{1,\mathcal{B}}) \right\} && (54) \\ &= n \left\{ |\mathcal{G}| \cdot |\mathcal{F}| \cdot \log_2 P + \Theta(1) \right\}, && (55) \end{aligned}$$

where the last step uses the fact that

$$h(Y^n_{\mathcal{F}_j}) \le \sum_{k \in \mathcal{F}_j} h(Y^n_k) \le n\{|\mathcal{F}| \log_2 P + O(1)\}, \tag{56}$$

and

$$h(Y^n_{\mathcal{F}_j} | M, X^n_{1,\mathcal{B}}) = h(Y^n_{\mathcal{F}_j} | X^n_{1,\mathcal{F}_j}, X^n_{2,\mathcal{F}_j}) = n \cdot O(1). \tag{57}$$

Dividing each side of (55) by $\log_2 P$ and taking the limit $P \to \infty$ yields (47).

Proof of Lemma 3: We use induction over the variable i to establish (52). For $i = 0$, note that $c_0 = 0$ and $\mathcal{V}_1 = \mathcal{G}$ and hence (52) is simply (49). This completes the proof for the base case.

For the induction step, we assume that (52) holds for some $t = i$, we need to show that (52) also holds for $t = i + 1$, i.e.,

$$\begin{aligned} & (i+1) \cdot n(R_{s,1} - \delta_n) + c_{i+1} \cdot n(R_{s,2} - \delta_n) \le \\ & \qquad \sum_{j=1}^{i+1} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) \end{aligned} \tag{58}$$

holds. For our proof we separately consider the cases when $|\mathcal{F}| \le |\mathcal{V}_i|$ and when $|\mathcal{V}_i| < |\mathcal{F}|$ holds.

When $|\mathcal{F}| \le |\mathcal{V}_i|$, from Definition 1

$$\mathcal{F}_{i+1} \subseteq \mathcal{V}_i, \qquad \mathcal{V}_{i+1} = \mathcal{V}_i \backslash \mathcal{F}_{i+1}, \qquad c_{i+1} = c_i \tag{59}$$

holds. Then (58) follows by combining (52) with (48) as we show below. Note that

$$\begin{aligned} & I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_i}) = I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{F}_{i+1}} | Y^n_{\mathcal{V}_i \backslash \mathcal{F}_{i+1}}) \\ & \qquad + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (60) \\ & \le I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}, Y^n_{\mathcal{V}_i \backslash \mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (61) \\ & \le I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}, X^n_{1,\mathcal{V}_i \backslash \mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (62) \\ & \le I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}, X^n_{1,\mathcal{G} \backslash \mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (63) \\ & = I(M, X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (64) \end{aligned}$$

where (60) follows from the chain rule of the mutual information and the definition of $\mathcal{V}_{i+1}$ in (59), while (62) follows from the Markov condition

$$Y^n_{\mathcal{V}_i \backslash \mathcal{F}_{i+1}} \leftrightarrow (X^n_{1,\mathcal{V}_i \backslash \mathcal{F}_{i+1}}, X^n_{2,\mathcal{V}_i \backslash \mathcal{F}_{i+1}}) \leftrightarrow (M, Y^n_{\mathcal{F}_{i+1}}, X^n_{1,\mathcal{B}\backslash\mathcal{G}}) \tag{65}$$

and the fact that $M = (X^n_{2,\mathcal{B}\cup\mathcal{C}}, Y^n_{1,\mathcal{A}})$ already includes $X^n_{2,\mathcal{V}_i \backslash \mathcal{F}_{i+1}}$, (63) follows from the fact that $\mathcal{V}_i \subseteq \mathcal{G}$, while (64) follows from the fact that $\{\mathcal{B}\backslash\mathcal{G}\} \cup \{\mathcal{G}\backslash\mathcal{F}_{i+1}\} = \{\mathcal{B}\backslash\mathcal{F}_{i+1}\}$.

Substituting (64) into the last term in (52) we get

$$\begin{aligned} & i \cdot n(R_{s,1} - \delta_n) + c_i \cdot n(R_{s,2} - \delta_n) \\ & \le \sum_{j=1}^{i} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_i}) \\ & \le \sum_{j=1}^{i} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) \\ & \qquad + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}). \end{aligned} \tag{66}$$

Finally combining (66) with (48) and using $c_{i+1} = c_i$ (c.f. (59)) we have

$$\begin{aligned} & (i+1) \cdot n(R_{s,1} - \delta_n) + c_{i+1} \cdot n(R_{s,2} - \delta_n) \\ & \le \sum_{j=1}^{i} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}}) \\ & \qquad + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) \\ & \qquad + I(X^n_{1,\mathcal{F}_{i+1}}; Y^n_{\mathcal{F}_{i+1}} | M, X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}}) && (67) \\ & = \sum_{j=1}^{i+1} I(M, X^n_{1,\mathcal{B}}; Y^n_{\mathcal{F}_j}) + I(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{V}_{i+1}}) && (68) \end{aligned}$$

as required.

When $|\mathcal{F}| > |\mathcal{V}_i|$, as stated in Definition 1 we introduce $\mathcal{H}_{i+1} = \{1, 2, \ldots, |\mathcal{F}| - |\mathcal{V}_i|\}$ and recall that

$$\mathcal{F}_{i+1} = \mathcal{V}_i \cup \mathcal{H}_{i+1}, \qquad \mathcal{V}_{i+1} = \mathcal{G}\backslash\mathcal{H}_{i+1}, \qquad c_{i+1} = c_i + 1 \tag{69}$$

holds. From (49) and (58) we have that

$$\begin{aligned}
&i\cdot n(R_{s,1}-\delta_n)+(c_i+1)\cdot n(R_{s,2}-\delta_n)\\
&=\sum_{j=1}^{i} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_i})\\
&\qquad+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{G}}) \qquad (70)\\
&=\sum_{j=1}^{i} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_i})\\
&\qquad+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{G}\backslash\mathcal{H}_{i+1}})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_{i+1}}) \qquad (71)
\end{aligned}$$

As we will show subsequently,

$$\begin{aligned}
&I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_i})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{G}\backslash\mathcal{H}_{i+1}})\\
&\qquad\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{F}_{i+1}}). \qquad (72)
\end{aligned}$$

Combining (48), (71) and (72) and using $c_{i+1}=c_i+1$ we get that

$$\begin{aligned}
&(i+1)\cdot n(R_{s,1}-\delta_n)+c_{i+1}\cdot n(R_{s,2}-\delta_n)\\
&\le\sum_{j=1}^{i} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_{i+1}})\\
&\quad+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{F}_{i+1}})+I(X^n_{\mathcal{F}_{i+1}};Y^n_{\mathcal{F}_{i+1}}|M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}}) \qquad (73)\\
&=\sum_{j=1}^{i} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_{i+1}})\\
&\qquad+I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_{i+1}}), \qquad (74)
\end{aligned}$$

which establishes (58).

It only remains to establish (72) which we do now. First, since $\mathcal{F}_{i+1}\subseteq\mathcal{G}$ it follows that $\{\mathcal{B}\backslash\mathcal{G}\}\subseteq\{\mathcal{B}\backslash\mathcal{F}_{i+1}\}$ and hence we bound the first term in the left hand side of (72) as

$$I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_i})\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{V}_i}). \qquad (75)$$

Next, since the set $\mathcal{H}_{i+1}=\{1,\ldots,|\mathcal{F}|-|\mathcal{V}_i|\}$ constitutes the first $|\mathcal{F}|-|\mathcal{V}_i|$ elements of $\mathcal{G}$ and $\mathcal{V}_i=\{|\mathcal{G}|-|\mathcal{V}_i|+1,\ldots,|\mathcal{G}|\}$ constitutes the last $|\mathcal{V}_i|$ elements of $\mathcal{G}$ and $|\mathcal{F}|\le|\mathcal{G}|$ we have that

$$\begin{aligned}
&\{\mathcal{G}\backslash\mathcal{H}_{i+1}\}=\{|\mathcal{F}|-|\mathcal{V}_i|+1,\ldots,|\mathcal{G}|\}\\
&=\{|\mathcal{F}|-|\mathcal{V}_i|+1,\ldots,|\mathcal{G}|-|\mathcal{V}_i|\}\cup\{|\mathcal{G}|-|\mathcal{V}_i|+1,\ldots,|\mathcal{G}|\}\\
&=\{\mathcal{G}\backslash(\mathcal{H}_{i+1}\cup\mathcal{V}_i)\}\cup\mathcal{V}_i\\
&=\{\mathcal{G}\backslash\mathcal{F}_{i+1}\}\cup\mathcal{V}_i \qquad (76)
\end{aligned}$$

where the last relation follows from the definition of $\mathcal{F}_{i+1}$ (c.f. (69)). Using (76) we can bound the second term in (72) as follows.

$$\begin{aligned}
&I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{G}\backslash\mathcal{H}_{i+1}})\\
&=I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{G}\backslash\mathcal{F}_{i+1}},Y^n_{\mathcal{V}_i}) \qquad (77)\\
&\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}},Y^n_{G\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{V}_i}) \qquad (78)\\
&\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}},X^n_{1,\mathcal{G}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{V}_i}) \qquad (79)\\
&\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{V}_i}), \qquad (80)
\end{aligned}$$

where in (79), we use the Markov relation

$$Y^n_{\mathcal{G}\backslash\mathcal{F}_{i+1}}\leftrightarrow(X^n_{1,\mathcal{G}\backslash\mathcal{F}_{i+1}},X^n_{2,\mathcal{G}\backslash\mathcal{F}_{i+1}})\leftrightarrow(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}},Y^n_{\mathcal{F}_{i+1}}) \qquad (81)$$

and the fact that $M=(X^n_{2,\mathcal{B}\cup\mathcal{C}},Y^n_{1,\mathcal{A}})$ already contains $X^n_{2,\mathcal{G}\backslash\mathcal{F}_{i+1}}$. Combining (75) and (80) gives

$$\begin{aligned}
&I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{V}_i})+I(M,X^n_{1,\mathcal{B}\backslash\mathcal{G}};Y^n_{\mathcal{H}_{i+1}}|Y^n_{\mathcal{G}\backslash\mathcal{H}_{i+1}})\\
&\qquad\le I(M,X^n_{1,\mathcal{B}\backslash\mathcal{F}_{i+1}};Y^n_{\mathcal{F}_{i+1}}), \qquad (82)
\end{aligned}$$

thus establishing (72).

This completes the proof.

D. Converse: $\min(|\mathcal{A}|,|\mathcal{C}|)\le N_E\le\max(|\mathcal{A}|,|\mathcal{C}|)$

We assume without loss of generality that $|\mathcal{C}|\ge|\mathcal{A}|$ and as before let $\mathcal{E}_k$ be the set of links such that an eavesdropper is monitoring for message W_k. Since $|\mathcal{E}_1|=|\mathcal{E}_2|=N_E$ and $|\mathcal{A}|\le N_E\le|\mathcal{C}|$ holds, we select the sets such that the relations $\mathcal{A}\subseteq\mathcal{E}_1\subseteq\mathcal{A}\cup\mathcal{B}$ and $\mathcal{C}\supseteq\mathcal{E}_2$ are both satisfied. Define $\mathcal{F}=\mathcal{B}\backslash\mathcal{E}_1$ and note that $|\mathcal{F}|=|\mathcal{A}|+|\mathcal{B}|-N_E$.

Theorem 1 reduces to the following region :

$$0\le d_1\le|\mathcal{F}| \qquad (83)$$
$$0\le d_2\le|\mathcal{B}|+|\mathcal{C}|-N_E \qquad (84)$$
$$|\mathcal{B}|d_1+|\mathcal{F}|d_2\le(|\mathcal{B}|+|\mathcal{C}|-N_E)\times|F| \qquad (85)$$

Since (83) and (84) directly follow from the single user case [17], we only need to establish (85). As in earlier cases we begin by establishing the following bounds on the rate pair $(R_{s,1},R_{s,2})$:

$$n(R_{s,1}-\delta_n)\le I(X^n_{1,\mathcal{F}};Y^n_{\mathcal{F}}|M,X^n_{1,\mathcal{B}\backslash\mathcal{F}}) \qquad (86)$$
$$n(R_{s,2}-\delta_n)\le I\left(M;Y^n_{\mathcal{B}}\right)+I\left(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2};Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) \qquad (87)$$

where $M=\left(X^n_{2,\mathcal{B}\cup\mathcal{C}},Y^n_{1,\mathcal{A}}\right)$.

Proof: The proof for (86) is identical to (48) in Lemma 2 since the proof does not depend on the choice of $\mathcal{E}_2$. The proof for (87) is identical to (42) in Lemma 1. ■

To establish (83)-(85), note that by defining

$$R'_{s,2}=R_{s,2}-\frac{1}{n}I\left(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2};Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right), \qquad (88)$$

we have from (87) that

$$n(R'_{s,2}-\delta_n)\le I\left(M;Y^n_{\mathcal{B}}\right) \qquad (89)$$

and the bounds on $R_{s,1}$ and $R'_{s,2}$ in (86) and (89) are identical to the bounds (48) and (49) in Lemma 2 with $\mathcal{G}=\mathcal{B}$. Applying Lemma 3 to $R_{s,1}$ and $R'_{s,2}$ for each $i=0,1,\cdots,|\mathcal{G}|$, it follows that

$$\begin{aligned}
&i\cdot n(R_{s,1}-\delta_n)+c_i\cdot n(R'_{s,2}-\delta_n)\\
&\le\sum_{j=1}^{i} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})+I(M,Y^n_{\mathcal{V}_i}). \qquad (90)
\end{aligned}$$

where the sets $\mathcal{V}_i$, $\mathcal{F}_i$ and the sequence c_i are as in Definition 1. Substituting (89) into (90) and evaluating the bound for $i=|\mathcal{B}|$

we have that

$$|\mathcal{B}|n(R_{s,1}-\delta_n)+|\mathcal{F}|n(R_{s,2}-\delta_n) \leq |\mathcal{F}|I(X^n_{2,\mathcal{C}\setminus\mathcal{E}_2};Y^n_{\mathcal{C}\setminus\mathcal{E}_2}) + \sum_{j=1}^{|\mathcal{B}|} I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j}). \quad (91)$$

Finally substituting

$$d\left(\frac{1}{n}I(M,X^n_{2,\mathcal{C}\setminus\mathcal{E}_2};Y^n_{\mathcal{C}\setminus\mathcal{E}_2}\right) \leq |\mathcal{C}|-N_E \quad (92)$$

$$d\left(\frac{1}{n}I(M,X^n_{1,\mathcal{B}};Y^n_{\mathcal{F}_j})\right) \leq |\mathcal{F}|, \quad (93)$$

in (91) we obtain (85). ■

VI. General MIMO-MAC

The result for the general MIMO case (1) follows by a transformation that reduces the model to the case of parallel independent channels in the previous section while preserving the secrecy degrees of freedom region. As we discuss next, this transformation involves the generalized singular value decomposition (GSVD) [28] and a channel enhancement argument. For an analogous application of GSVD to broadcast channels see e.g., [18], [29], [30].

A. GSVD Transformation

Theorem 2: [28] Given a pair of matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ such that the rank of $\mathbf{H}_i$ is r_i, $i=1,2$, and the rank of $[\ \mathbf{H}_1 \mid \mathbf{H}_2\ ]$ is r_0, there exists unitary matrices $\mathbf{U}_1,\mathbf{U}_2,\mathbf{W},\mathbf{Q}$ and nonsingular upper triangular matrix $\mathbf{R}$ such that for $s=r_1+r_2-r_0$, $\tilde{r}_1=r_1-s$, $\tilde{r}_2=r_2-s$,

$$\mathbf{U}_1^H\mathbf{H}_1^H\mathbf{Q}=\mathbf{\Sigma}_{1(N_{T_1}\times r_0)}\left[\mathbf{W}^H\mathbf{R}_{(r_0\times r_0)},\mathbf{0}\right]_{(r_0\times N_R)} \quad (94)$$

$$\mathbf{U}_2^H\mathbf{H}_2^H\mathbf{Q}=\mathbf{\Sigma}_{2(N_{T_2}\times r_0)}\left[\mathbf{W}^H\mathbf{R}_{(r_0\times r_0)},\mathbf{0}\right]_{(r_0\times N_R)} \quad (95)$$

$$\mathbf{\Sigma}_1=\begin{bmatrix}\mathbf{I}_{1(\tilde{r}_1\times\tilde{r}_1)} & & \\ & \mathbf{S}_{1(s\times s)} & \\ & & \mathbf{O}_{1((N_{T_1}-\tilde{r}_1-s)\times\tilde{r}_2)}\end{bmatrix} \quad (96)$$

$$\mathbf{\Sigma}_2=\begin{bmatrix}\mathbf{O}_{2((N_{T_2}-\tilde{r}_2-s)\times\tilde{r}_1)} & & \\ & \mathbf{S}_{2(s\times s)} & \\ & & \mathbf{I}_{2(\tilde{r}_2\times\tilde{r}_2)}\end{bmatrix} \quad (97)$$

where $\mathbf{I}_i, i=1,2$ are $\tilde{r}_i\times\tilde{r}_i$ identity matrices, $\mathbf{O}_i, i=1,2$ are zero matrices, and $\mathbf{S}_i, i=1,2$ are $s\times s$ diagonal matrices with positive real elements on the diagonal line that satisfy $\mathbf{S}_1^2+\mathbf{S}_2^2=\mathbf{I}_s$, and $\tilde{r}_1+s+\tilde{r}_2=r_0$. For clarity, the dimension of each matrix is shown in the parenthesis in the subscript. $\mathbf{I}_1$ has the same number of columns as $\mathbf{O}_2$. $\mathbf{I}_2$ has the same number of columns $\mathbf{O}_1$. However, $\mathbf{O}_i, i=1,2$ are not necessarily square matrices and can be empty, i.e., having zero number of rows.

For convenience in notation we define $\mathbf{A}=\mathbf{W}^H\mathbf{R}$ and observe that $\mathbf{A}$ is a square and non-singular matrix. Then from Theorem 2, we have:

$$\mathbf{Q}^H\mathbf{H}_t\mathbf{U}_t=\begin{bmatrix}\mathbf{A}^H\\ \mathbf{0}\end{bmatrix}\mathbf{\Sigma}_t^H, t=1,2. \quad (98)$$

Without loss of generality, we can cancel $\mathbf{Q}$ and $\mathbf{U}_t$ and rewrite (1) as:

$$\mathbf{Y}=\begin{bmatrix}\mathbf{A}^H_{r_0\times r_0}\\ \mathbf{0}_{(N_R-r_0)\times r_0}\end{bmatrix}_{N_R\times r_0}\mathbf{\Sigma}_1^H\mathbf{X}_1 + \begin{bmatrix}\mathbf{A}^H_{r_0\times r_0}\\ \mathbf{0}_{(N_R-r_0)\times r_0}\end{bmatrix}_{N_R\times r_0}\mathbf{\Sigma}_2^H\mathbf{X}_2+\mathbf{Z}. \quad (99)$$

Since $\mathbf{Q}$ and $\mathbf{U}_t$ are unitary matrices, the components of $\mathbf{Z}$ are independent from each other and the power constraints of each transmitter remains the same as $\bar{P}_i, i=1,2$. Because the components of $\mathbf{Z}$ are independent, the intended receiver can discard the last N_R-r_0 components in $\mathbf{Y}$ without affecting the secrecy capacity region of this channel. This means that we only need to consider the case where $N_R=r_0$ and rewrite (1) as:

$$\mathbf{Y}=\mathbf{A}^H_{r_0\times r_0}(\mathbf{\Sigma}_1^H\mathbf{X}_1+\mathbf{\Sigma}_2^H\mathbf{X}_2)+\mathbf{Z}. \quad (100)$$

B. Converse

For establishing the converse, we further enhance the channel model in (100) to the following

$$\mathbf{Y}=\mathbf{\Sigma}_1^H\mathbf{X}_1+\mathbf{\Sigma}_2^H\mathbf{X}_2+\sigma_+\mathbf{Z}' \quad (101)$$

where $\sigma_+\leq 1$ is any sufficiently small constant such that, σ_+^2 times the maximal eigenvalue of $\mathbf{A}^H_{r_0\times r_0}\mathbf{A}_{r_0\times r_0}$, is smaller than 1 and $\mathbf{Z}'$ is a circularly symmetric unit-variance Gaussian noise vector.

To establish (101), note that we can express

$$\mathbf{Z}=\sigma_+\cdot\mathbf{A}^H\mathbf{Z}'+\mathbf{Z}'' \quad (102)$$

where $\mathbf{Z}''$ is a Gaussian random vector, independent of $\mathbf{Z}'$ and with a covariance matrix

$$\mathbf{I}_{r_0\times r_0}-\sigma_+^2\mathbf{A}^H_{r_0\times r_0}\mathbf{A}_{r_0\times r_0} \quad (103)$$

which is guaranteed to be positive semi-definite by our choice of σ_+. Upon substituting (102) into (100), we have

$$\mathbf{Y}=\mathbf{A}^H_{r_0\times r_0}\left(\mathbf{\Sigma}_1^H\mathbf{X}_1+\mathbf{\Sigma}_2^H\mathbf{X}_2+\sigma_+\mathbf{Z}'\right)+\mathbf{Z}''. \quad (104)$$

We consider an enhanced receiver that is revealed $\mathbf{Z}''$. Clearly this additional knowledge can only increase the rate and serves as an upper bound. It is also clear that since $\mathbf{Z}''$ is independent of $(\mathbf{X}_1,\mathbf{X}_2,\mathbf{Z}')$, it suffices to use this information to cancel $\mathbf{Z}''$ in (104) and then discard it. Furthermore since the matrix $\mathbf{A}$ is invertible, upon canceling it, we obtain (101).

We further enhance the receiver by replacing $\mathbf{\Sigma}_1^H$ and $\mathbf{\Sigma}_2^H$ with $\bar{\mathbf{\Sigma}}_1^H$ and $\bar{\mathbf{\Sigma}}_2^H$ so that the model reduces to

$$\mathbf{Y}=\bar{\mathbf{\Sigma}}_1^H\mathbf{X}_1+\bar{\mathbf{\Sigma}}_2^H\mathbf{X}_2+\sigma_+\mathbf{Z}' \quad (105)$$

where

$$\bar{\mathbf{\Sigma}}_1^H=\begin{bmatrix}\mathbf{I}_{\tilde{r}_1\times\tilde{r}_1} & & \\ & \mathbf{I}_{1(s\times s)} & \\ & & \mathbf{0}_{1\left(\tilde{r}_2\times\left(N_{T_1}-r_1\right)\right)}\end{bmatrix}_{r_0\times N_{T_1}} \quad (106)$$

$$\bar{\mathbf{\Sigma}}_2^H = \begin{bmatrix} \mathbf{0}_{2\left(\tilde{r}_1 \times \left(N_{T_2} - r_2\right)\right)} & & \\ & \mathbf{I}_{2(s \times s)} & \\ & & \mathbf{I}_{2(\tilde{r}_2 \times \tilde{r}_2)} \end{bmatrix}_{r_0 \times N_{T_2}} \tag{107}$$

are obtained by replacing each diagonal $\mathbf{S}_i$ by the identity matrix. The model (105) can only have a higher capacity, since each diagonal entry in $\mathbf{S}_i$ is between $(0, 1)$. We observe that in the resulting channel model is identical to (22)-(24)

$$|\mathcal{A}| = r_0 - r_2 \tag{108}$$

$$|\mathcal{B}| = s = r_1 + r_2 - r_0 \tag{109}$$

$$|\mathcal{C}| = r_0 - r_1 \tag{110}$$

except that the noise variance is reduced by a factor of σ_+^2. Since a fixed scaling in the noise power does not affect the secure-degrees of freedom, an outer bound on the s.d.o.f. for the parallel channel model (22)-(24) with $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ defined via (105), continues to be an outer bound on the s.d.o.f. region for the general MIMO-MAC channel.

Substituting (108)-(110) in the upper bounds in section V-B, V-C and V-D we establish the converse in Theorem 1.

C. Achievability

To establish the achievability for the general MIMO case we further use a suitable degradation mechanism to reduce the model (100) to

$$\mathbf{Y} = \mathbf{\Sigma}_1^H \mathbf{X}_1 + \mathbf{\Sigma}_2^H \mathbf{X}_2 + \sigma \mathbf{Z}'' \tag{111}$$

where $\sigma \geq 1$ is any sufficiently large constant such that, σ^2 times the minimum eigenvalue of $\mathbf{A}_{r_0 \times r_0}^H \mathbf{A}_{r_0 \times r_0}$, is greater than 1 and $\mathbf{Z}''$ is a circularly symmetric unit-variance Gaussian noise vector. Since $\mathbf{A}$ is non-singular we are guaranteed that all the singular values of $\mathbf{A}$ are non-zero and hence a $\sigma < \infty$ exists.

To establish (111), let $\mathbf{Z}'$ be a Gaussian noise vector with covariance

$$\sigma^2 \mathbf{A}_{r_0 \times r_0}^H \mathbf{A}_{r_0 \times r_0} - \mathbf{I}_{r_0 \times r_0} \tag{112}$$

independent of $\mathbf{Z}$ and consider a degraded version of (100)

$$\mathbf{Y} = \mathbf{A}_{r_0 \times r_0}^H \left(\mathbf{\Sigma}_1^H \mathbf{X}_1 + \mathbf{\Sigma}_2^H \mathbf{X}_2 \right) + \mathbf{Z} + \mathbf{Z}' \tag{113}$$

which can be simulated at the receiver by adding additional noise $\mathbf{Z}'$ to its output. Since $\mathbf{Z} + \mathbf{Z}' \sim \mathcal{CN}(0, \sigma^2 \mathbf{A}^H \mathbf{A})$, we can express $\mathbf{Z} + \mathbf{Z}' = \sigma \mathbf{A}^H \mathbf{Z}''$. Substituting into (113) and canceling the non-singular matrix $\mathbf{A}$, we arrive at (111).

Let $\bar{s} > 0$ denote the minimum element on the diagonals of $\mathbf{S}_1$ and $\mathbf{S}_2$ in (96) and (97) respectively. By appropriately scaling down the transmit powers on each of the sub-channels we can further reduce (104) to

$$\mathbf{Y} = \bar{\mathbf{\Sigma}}_1^H \mathbf{X}_1 + \bar{\mathbf{\Sigma}}_2^H \mathbf{X}_2 + \frac{\sigma}{\bar{s}} \mathbf{Z}'' \tag{114}$$

where $\bar{\mathbf{\Sigma}}_k$ are defined in (96) and (97) respectively. The model (114) is identical to the parallel channel model (22)-(24) with the size of sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in (108)-(110) and with a noise power that is larger by a factor of $\sigma^2/\bar{s}^2$. Since a constant factor in the noise power does not affect the secrecy degrees of freedom, the coding schemes described in section V-A achieves the lower bound in Theorem 1.

VII. Conclusion

In this work we have studied the two-transmitter Gaussian complex MIMO-MAC wiretap channel where the eavesdropper channel is arbitrarily varying and its state is known to the eavesdropper only, and the main channel is static and its state is known to all nodes. We have completely characterized the s.d.o.f. region for this channel for all possible antenna configurations. We have proved that this s.d.o.f. region can be achieved by a scheme that orthogonalizes the transmit signals of the two users at the intended receiver, in which each user achieves secrecy guarantee independently without cooperation from the other user. The converse was proved by carefully changing the set of signals available to the eavesdropper through an induction procedure in order to obtain an upper bound on a weighted-sum-rate expression.

As suggested by this work, the optimal strategy for a communication network where the eavesdropper channel is arbitrarily varying can potentially be very different from the case where the eavesdropper channel is fixed and its state is known to all terminals. This is also observed for example in the MIMO broadcast channel [18] and the two-way channel [31], [32]. Characterizing secure transmission limits for a broader class of communication models with this assumption is hence important and is left as future work.

Appendix A
Proof of Lemma 1

For $R_{s,1}$, from Fano's inequality, we have

$$\begin{aligned} n(R_{s,1} - \delta_n) &\leq I(W_1; Y_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}^n) - I(W_1; Y_{\mathcal{E}_1}^n) && (115) \\ &\leq I\left(W_1; Y_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}^n | Y_{\mathcal{E}_1}^n\right) && (116) \\ &\leq I\left(W_1; Y_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n | Y_{\mathcal{E}_1}^n\right) && (117) \\ &= I\left(W_1; Y_{\mathcal{A}\cup\mathcal{B}}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n | Y_{\mathcal{E}_1}^n\right) && (118) \end{aligned}$$

where the last step (118) relies on the fact that the additive noise at each receiver end of each sub-channel in Figure 5 is independent from each other and hence

$$Y_{\mathcal{C}}^n \to X_{2,\mathcal{C}}^n \to (W_1, Y_{\mathcal{A}\cup\mathcal{B}}^n, Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}}^n)$$

holds. Since $\left(X_{2,\mathcal{C}}^n, X_{2,\mathcal{B}}^n\right)$ is independent from W_1, and $\mathcal{E}_1 \subseteq \mathcal{A}$, (118) can be written as:

$$\begin{aligned} &I\left(W_1; Y_{\mathcal{A}\cup\mathcal{B}}^n | Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n\right) \\ =&I\left(W_1; Y_{(\mathcal{A}\backslash\mathcal{E}_1)\cup\mathcal{B}}^n | Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n\right) && (119) \\ =&I\left(W_1; Y_{\mathcal{A}\backslash\mathcal{E}_1}^n | Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n\right) + I\left(W_1; Y_{\mathcal{B}}^n | Y_{\mathcal{A}}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n\right) && (120) \end{aligned}$$

where the last step (120) follows from the fact $\mathcal{E}_1 \subseteq \mathcal{A}$ and hence $\mathcal{A} = (\mathcal{A}\backslash\mathcal{E}_1) \cup \mathcal{E}_1$. We separately bound each of the two terms above.

$$\begin{aligned} &I\left(W_1; Y_{\mathcal{A}\backslash\mathcal{E}_1}^n | Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n\right) \\ &\leq I\left(W_1, Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n; Y_{\mathcal{A}\backslash\mathcal{E}_1}^n\right) && (121) \\ &\leq I\left(W_1, Y_{\mathcal{E}_1}^n, X_{2,\mathcal{B}\cup\mathcal{C}}^n, X_{1,\mathcal{A}\backslash\mathcal{E}_1}^n; Y_{\mathcal{A}\backslash\mathcal{E}_1}^n\right) && (122) \\ &= I\left(X_{1,\mathcal{A}\backslash\mathcal{E}_1}^n; Y_{\mathcal{A}\backslash\mathcal{E}_1}^n\right) && (123) \end{aligned}$$

where the last step follows from the Markov chain relation $Y^n_{1,\mathcal{A}\backslash\mathcal{E}_1} \leftrightarrow X^n_{\mathcal{A}\backslash\mathcal{E}_1} \leftrightarrow (W_1, Y^n_{\mathcal{E}_1}, X^n_{2,\mathcal{B}\cup\mathcal{C}})$, We upper bound the second term in (120) as follows

$$\begin{aligned}&I\left(W_1; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right)\\ &\leq I\left(X^n_{1,\mathcal{A}\cup\mathcal{B}}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(124)\\ &= I\left(X^n_{1,\mathcal{B}}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right)\\ &\quad + I\left(X^n_{1,\mathcal{A}}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}, X^n_{1,\mathcal{B}}\right) &(125)\\ &= I\left(X^n_{1,\mathcal{B}}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(126)\end{aligned}$$

where we use the Markov relation $W_1 \leftrightarrow X^n_{1,\mathcal{A}\cup\mathcal{B}} \leftrightarrow (Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}})$ in step (124) and (126) follows from the fact Markov relation

$$Y^n_{\mathcal{B}} \leftrightarrow (X^n_{1,\mathcal{B}}, X^n_{2,\mathcal{B}}) \leftrightarrow (X^n_{2,\mathcal{C}}, Y^n_{\mathcal{A}}). \quad (127)$$

Note that (41) follows upon substituting (123) and (126) into (120).

For $R_{s,2}$, from Fano's inequality and the secrecy constraint, we have:

$$\begin{aligned}&n(R_{s,2}-\delta_n) \leq I(W_2; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}) - I(W_2; X^n_{2,\mathcal{E}_2}) &(128)\\ &\leq I\left(W_2; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}|X^n_{2,\mathcal{E}_2}\right) &(129)\\ &= I\left(W_2; Y^n_{\mathcal{B}\cup\mathcal{C}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}\right) &(130)\\ &= I\left(W_2; Y^n_{(\mathcal{C}\backslash\mathcal{E}_2)\cup\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}\right) &(131)\\ &= I\left(W_2; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}\right) + I\left(W_2; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) &(132)\end{aligned}$$

where (130) follows from the fact that $Y^n_{\mathcal{A}}$ is independent of $(W_2, X^n_{2,\mathcal{B}\cup\mathcal{C}})$ and (131) follows from the fact that $Y^n_{\mathcal{E}_2} \to X^n_{2,\mathcal{E}_2} \to (Y^n_{\mathcal{B}\cup\mathcal{C}\backslash\mathcal{E}_2}, W_2, Y^n_{\mathcal{A}})$ holds. We separately bound each term in (132).

$$\begin{aligned}&I\left(W_2; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}\right) \leq I\left(W_2, Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) &(133)\\ &\leq I\left(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2}, W_2, Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) &(134)\\ &= I\left(X^n_{2,\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right), &(135)\end{aligned}$$

where the justification for establishing (135) is identical to (123) and hence omitted. We finally bound the second term in (132).

$$\begin{aligned}&I\left(W_2; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) &(136)\\ &\leq I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{C}\backslash\mathcal{E}_2}\right) &(137)\\ &\leq I\left(Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}, X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{\mathcal{B}}\right) &(138)\\ &= I\left(Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}, X^n_{2,\mathcal{E}_2}; Y^n_{\mathcal{B}}\right)\\ &\quad + I\left(Y^n_{\mathcal{C}\backslash\mathcal{E}_2}; Y^n_{\mathcal{B}}|Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}, X^n_{2,\mathcal{E}_2}\right) &(139)\\ &= I\left(Y^n_{\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{B}}\right) &(140)\end{aligned}$$

where the justification for arriving at (140) is similar to (126) and hence omitted.

Substituting (135) and (140) into (132) we establish (42).

Appendix B
Proof of Lemma 2

Assume the eavesdropper monitors $Y^n_{\mathcal{A}}$ and $X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}$ for W_1. Then for $R_{s,1}$, from Fano's inequality, we have:

$$\begin{aligned}&n(R_{s,1}-\delta_n)\\ &\leq I\left(W_1; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}\right) - I\left(W_1; Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}\right) &(141)\\ &\leq I\left(W_1; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}\right) &(142)\\ &= I\left(W_1; Y^n_{\mathcal{B}\cup\mathcal{C}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}\right) &(143)\\ &\leq I\left(W_1; Y^n_{\mathcal{B}\cup\mathcal{C}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}\right) &(144)\\ &= I\left(W_1; Y^n_{\mathcal{B}\cup\mathcal{C}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(145)\\ &= I\left(W_1; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(146)\\ &= I\left(W_1; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(147)\end{aligned}$$

where (145) follows from the fact that $X^n_{2,\mathcal{B}\cup\mathcal{C}}$ is independent of $(W_1, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_1\backslash\mathcal{A}})$. while (146) follows from the fact that since the noise across the channels is independent the Markov condition

$$(Y^n_{\mathcal{E}_1\backslash\mathcal{A}}, Y^n_{\mathcal{C}}) \leftrightarrow (X^n_{1,\mathcal{E}_1\backslash\mathcal{A}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}) \leftrightarrow (W_1, Y^n_{B\backslash\mathcal{E}_1}, Y^n_{\mathcal{A}})$$

holds and furthermore we have defined $\mathcal{F} = \mathcal{B}\backslash\mathcal{E}_1$.

Since the channel noise is independent of the message, $W_1 \leftrightarrow X^n_{1,\mathcal{A}\cup\mathcal{B}} \leftrightarrow (Y^n_{\mathcal{F}\cup\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}})$ holds. Hence

$$\begin{aligned}&I\left(W_1; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(148)\\ &\leq I\left(X^n_{1,\mathcal{A}\cup\mathcal{B}}; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(149)\\ &= I\left(X^n_{1,\mathcal{F}}; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right)\\ &\quad + I\left(X^n_{1,\mathcal{A}\cup\mathcal{B}\backslash\mathcal{F}}; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(150)\\ &= I\left(X^n_{1,\mathcal{F}}; Y^n_{\mathcal{F}}|Y^n_{\mathcal{A}}, X^n_{1,\mathcal{B}\backslash\mathcal{F}}, X^n_{2,\mathcal{B}\cup\mathcal{C}}\right) &(151)\end{aligned}$$

where the last step uses the fact that the second term in (150) involves conditioning on $(X^n_{1,\mathcal{F}}, X^n_{2,\mathcal{F}})$ and hence is zero. This establishes (48).

For $R_{s,2}$, we assume the eavesdropper is monitoring $X^n_{2,\mathcal{C}}, X^n_{2,\mathcal{E}_2\backslash\mathcal{C}}$ for W_2. Using Fano's inequality and the secrecy constraint, we have:

$$\begin{aligned}&n(R_{s,2}-\delta_n) \leq I\left(W_2; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}\right) - I\left(W_2; X^n_{2,\mathcal{E}_2}\right) &(152)\\ &\leq I\left(W_2; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}|X^n_{2,\mathcal{E}_2}\right) &(153)\\ &\leq I\left(W_2; Y^n_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}|X^n_{2,\mathcal{E}_2}\right) &(154)\\ &= I\left(W_2; Y^n_{\mathcal{B}\cup\mathcal{C}}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right) &(155)\\ &\leq I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{B}\cup\mathcal{C}}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right) &(156)\\ &= I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{G}\cup\mathcal{E}_2}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right) &(157)\\ &= I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{G}}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right)\\ &\quad + I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{E}_2}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}\cup\mathcal{G}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right) &(158)\\ &= I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}; Y^n_{\mathcal{G}}|X^n_{2,\mathcal{E}_2}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}\right) &(159)\\ &\leq I\left(X^n_{2,\mathcal{B}\cup\mathcal{C}}, Y^n_{\mathcal{A}}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}; Y^n_{\mathcal{G}}\right) &(160)\end{aligned}$$

$$\leq I\left(M, X^n_{1,\mathcal{B}\backslash\mathcal{G}}; Y^n_{\mathcal{G}}\right) \tag{161}$$

where (155) follows from the fact that $(X^n_{1,\mathcal{E}_2\cap\mathcal{B}}, Y^n_{\mathcal{A}})$ are the transmitted signals from user 1 and independent of $(W_2, X^n_{2,\mathcal{E}_2})$ and (157) follows from the fact that $\mathcal{C} \subseteq \mathcal{E}_2 \subseteq \mathcal{B}\cup\mathcal{C}$ and $\mathcal{G} = \mathcal{B}\backslash\mathcal{E}_2$ and hence $\mathcal{E}_2\cup\mathcal{G} = \mathcal{B}\cup\mathcal{C}$ holds. Eq. (159) follows from the fact that since the noise on each channel is Markov, we have $Y^n_{\mathcal{E}_2} \leftrightarrow (X^n_{2,\mathcal{E}_2}, X^n_{1,\mathcal{E}_2\cap\mathcal{B}}) \leftrightarrow (Y^n_{\mathcal{A}\cup\mathcal{G}}, X^n_{\mathcal{B}\cup\mathcal{C}})$ and hence the second term in (158) is zero.

Hence we have proved Lemma 2.

References

[1] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, September 1949.

[2] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang, H. Poor, and S. Shamai Shitz, "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2009.

[4] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge Univ Pr, 2011.

[5] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. McLaughlin, and J. Merolla, "On the application of LDPC codes to a novel wiretap channel inspired by quantum key distribution," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, August 2005.

[6] M. Bellare and S. Tessaro, "Polynomial-Time, Semantically-Secure Encryption Achieving the Secrecy Capacity," 2012, available online at http://arxiv.org/abs/1201.3160.

[7] X. Li, J. Hwu, and E. P. Ratazzi, "Using Antenna Array Redundancy and Channel Diversity for Secure Wireless Transmissions," *Journal of Communications*, vol. 2, no. 3, pp. 24–32, 2007.

[8] S. Goel and R. Negi, "Guaranteeing Secrecy using Artificial Noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, June 2008.

[9] A. Khisti, G. W. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *IEEE International Symposium on Information Theory*, June 2007.

[10] A. Khisti and G. Wornell, "Secure Transmission with Multiple Antennas-I: The MISOME Wiretap Channel," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3088–3104, July 2010.

[11] ——, "Secure Transmission with Multiple Antennas-II: The MIMOME Wiretap Channel," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, November 2010.

[12] R. Liu, T. Liu, and H. V. Poor, "Multiple-input Multiple-output Gaussian Broadcast Channels with Confidential Messages," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4215–4227, September 2010.

[13] E. Ekrem and S. Ulukus, "The Secrecy Capacity Region of the Gaussian MIMO Multi-receiver Wiretap Channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2083–2114, April 2011.

[14] M. Kobayashi, Y. Liang, S. Shamai, and M. Debbah, "On the Compound MIMO Broadcast Channels with Confidential Messages," in *IEEE International Symposium on Information Theory*, June 2009.

[15] A. Mukherjee and A. L. Swindlehurst, "Robust Beamforming for Security in MIMO Wiretap Channels with Imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 351–361, 2011.

[16] E. MolavianJazi, "Secure Communication Over Arbitrarily Varying Wiretap Channels," *Master Thesis*, December 2009, available online at http://etd.nd.edu/ETD-db/theses/available/etd-12112009-112419/unrestricted/MolavianJaziE122009.pdf.

[17] X. He and A. Yener, "MIMO Wiretap Channels with Arbitrarily Varying Eavesdropper Channel States," submitted to the IEEE Transactions on Information Theory, July, 2010, in revision, available online at http://arxiv.org/abs/1007.4801.

[18] X. He, A. Khisti, and A. Yener, "MIMO Broadcast Channel with Arbitrarily Varying Eavesdropper Channel: Secrecy Degrees of Freedom," in *IEEE Global Telecommunication Conference*, December 2011.

[19] E. Tekin and A. Yener, " The General Gaussian Multiple Access and Two-Way Wire-Tap Channels: Achievable Rates and Cooperative Jamming," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2735–2751, June 2008.

[20] Y. Liang and H. V. Poor, "Multiple Access Channels With Confidential Messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[21] E. Ekrem and S. Ulukus, "On the Secrecy of Multiple Access Wiretap Channel," in *46th Allerton Conference on Communication, Control, and Computing*, September 2008.

[22] D. S. Papailiopoulos and A. G. Dimakis, "Distributed Storage Codes Meet Multiple-Access Wiretap Channels," in *48th Allerton Conference on Communication, Control and Computing*, September 2010.

[23] X. He and A. Yener, "Providing Secrecy With Structured Codes: Tools and Applications to Gaussian Two-user Channels," submitted to IEEE Transactions on Information Theory, July, 2009, in revision, available online at http://arxiv.org/abs/0907.5388.

[24] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "On the Secure Degrees-of-Freedom of the Multiple-Access-Channel," in *IEEE International Symposium on Information Theory*, June 2010.

[25] X. He, "Cooperation and Information Theoretic Security in Wireless Networks," *PhD Dissertation*, August 2009, available online at http://etda.libraries.psu.edu/theses/approved/WorldWideIndex/ETD-5342/index.html.

[26] A. Khisti, "Interference Alignment for the Multi-Antenna Compound Wiretap Channel," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2967–2993, May 2011.

[27] D. N. C. Tse and S. V. Hanly, "Multiaccess Fading Channels. I. Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.

[28] C. C. Paige and M. A. Saunders, "Towards a generalized singular value decomposition," *SIAM Journal on Numerical Analysis*, vol. 18, no. 3, pp. 398–405, 1981.

[29] A. Khisti, D. Silva, and F. Kschischang, "Secure Broadcast Codes over linear determintic channels," in *IEEE International Symposium on Information Theory*, May 2010.

[30] E. Ekrem and S. Ulukus, "Degrees of Freedom Region of the Gaussian MIMO Broadcast Channel with Common and Private Messages," in *IEEE Global Telecommunication Conference*, December 2010.

[31] X. He and A. Yener, "Secrecy When the Eavesdropper Controls its Channel States," in *IEEE International Symposium on Information Theory*, July 2011.

[32] ——, "Gaussian Two-way Wiretap Channel with an Arbitrarily Varying Eavesdropper," in *IEEE Global Telecommunication Conference, Workshop on Physical Layer Security*, December 2011.